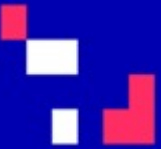




claudette.eui.eu



Text analytics in the legal domain:

the case of contracts and privacy policies in Claudette

Francesca Lagioia, PhD







francesca.lagioia@eui.eu, francesca.lagioia@unibo.it



The Claudette Team

Law Dept@eui

Cirsfid@unibo

 Hans-Wolfgang Micklits	 Kasper Drazewski	 Giacomo Tagiuri	 Agnieszka Jablonowska	 Francesca Lagioia	 Giovanni Sartor
---	---	---	--	--	--

 Przemyslaw Palka
--

Law School@yale

 Marco Coppi

DISMI@unimore

 Paolo Torroni	 Federico Ruggeri	 Andrea Galassi
---	--	--

DISI@unibo

 Ruta Liepina
--

Law Dept@Maastricht



claudette.eui.eu

NEW COUNTOURS OF AI AND LAW

Recently, the popular perception of AI is that of **something** at the service of businesses

AI is **currently affecting** consumers' ...

- privacy
- autonomy
- economic interests
- behaviour
- access to goods and services
- social exclusion

That does not have to be the case!
AI can unlock **consumer-empowering technologies**



How to empower consumers?

- Protection against unwanted monitoring (GDPR)
- Support in detecting unfair use of AI
- Control commercial practice fairness

“An opposing exercise of power is the principal solvent of economic power, the basic defense against its exercise in economic affairs”. Ken Galbraith

In the AI era an effective countervailing power needs to be supported by AI.

CLAUDETTE.eui.eu

Automatically detect **potentially** unfair clauses in Terms of Services and Privacy Policies

- Consumers agree but don't read
- NGOs have competence to control but lack resources
- Business keeps using unlawful clauses

Terms of Service (ToS): The Training Set

The ToS Corpus

WHERE DID WE START?

... 50 ToS (manually annotated)...

7,090 sentences, 787 of which (11.1%) were labeled as positive, thus containing a potentially unfair clause.

WHERE ARE WE NOW?

... 100 ToS (manually annotated)...



Part 1: Unfair Contract Terms Law and Practice

Directive 93/13 art 3.1:

A contractual term which has **not** been **individually negotiated** shall be regarded as **unfair** if, contrary to the requirement of good faith, it causes a **significant imbalance** in the parties' rights and obligations arising under the contract, to the detriment of the consumer.

Bottom-line: there are some types of clauses that traders are prohibited from using in the contracts.

8 unfairness categories (Art. 3 of Directive 93/13)

Type of clause	Symbol	# clauses (50 Tos)	#documents (50 Tos)
Arbitration	<a>	44	28
Unilateral change	<ch>	188	49
Content removal	<c>	118	45
Jurisdiction	<j>	68	40
Choice of law	<law>	70	47
Limitation of liability	<ltld>	296	49
Unilateral termination	<ter>	236	48
Consent by using	<use>	117	48
Privacy included	<pinc>		

1) clearly fair; 2) potentially unfair; 3) clearly unfair

claudette.eui.eu

Consent by using Clause

If a clause states that the consumer is bound by the terms of service simply by visiting the website or by downloading the app, or by using the service: **potentially unfair**

A **potentially unfair** consent by using clause (Airbnb):

```
<use2>By accessing or using the Airbnb Platform, you agree to comply with and be bound by these Terms of Service.</use2>
```

A **potentially unfair** consent by using clause (Facebook):

```
<use2>By using or accessing the Facebook Services, you agree to this Statement, as updated from time to time in accordance with Section 13 below.</use2>
```

Jurisdiction Clause

Where a dispute will be adjudicated?

If giving consumers a right to bring disputes in their place of residence: **clearly fair**

If stating that any judicial proceeding takes a residence away (i.e. in a different city, different country): **clearly unfair**

A clearly unfair jurisdiction clause (Dropbox):

```
<j3> You and Dropbox agree that any judicial proceeding to resolve claims relating to these Terms or the Services will be brought in the federal or state courts of San Francisco County, California, subject to the mandatory arbitration provisions below. Both you and Dropbox consent to venue and personal jurisdiction in such courts.</j3>
```

Limitation of Liability

For what actions/events the provider claims they will not be liable?

If stating that the provider may be liable: **clearly fair**

If stating that the provider will never be liable for any action taken by other people// damages incurred by the computer because of malware // When contains a blanket phrase like “to the fullest extent permissible by law”: **potentially unfair**

If stating that the provider will never be liable for physical injuries (health/life)// gross negligence// intentional damage: **clearly unfair**

Limitation of Liability

For what actions/events the provider claims they will not be liable?

A **fair liability** clause (World of Warcraft):

```
<ltd1>Blizzard Entertainment is liable in accordance with statutory law (i) in case of intentional breach, (ii) in case of gross negligence, (iii) for damages arising as result of any injury to life, limb or health or (iv) under any applicable product liability act.</ltd1>
```

A **potentially unfair** limitation of liability clause (9gag):

```
<ltd2>You agree that neither 9GAG, Inc nor the Site will be liable in any event to you or any other party for any suspension, modification, discontinuance or lack of availability of the Site, the service, your Subscriber Content or other Content.</ ltd2>
```

Limitation of Liability

For what actions/events the provider claims they will not be liable?

A **potentially unfair** limitation of liability clause (Truecaller):

<1td2>To the maximum extent permitted by applicable law, you expressly agree that truecaller shall in no event be liable for any direct, indirect, special, incidental, consequential or exemplary damages, including but not limited to damages for loss of profits, data and goodwill, arising out of the use or inability to use the services or the content, even if advised of the possibility of such damages</1td2>

Limitation of Liability

For what actions/events the provider claims they will not be liable?

A **clearly unfair** limitation of liability clause (Rovio):

In no event will Rovio, Rovio's affiliates, Rovio's licensors or channel partners be liable for special, incidental or consequential damages resulting from possession, access, use or malfunction of the Rovio services, [...] and, to the extent permitted by law, damages for **personal injuries**, [...] whether or not Rovio, Rovio's licensors or channel partners have been advised of the possibility of such damages.

Privacy Included

Whenever a clause states (or it might be possible to assume) that the consumer consents to the privacy policy simply by using the service: **potentially unfair**

A potentially unfair clause (**DeviantArt**):

<use2>By using our Service, you agree to be bound by Section I of these Terms ("General Terms"), which contains provisions applicable to all users of our Service, including visitors to the DeviantArt website (the "Site").</use2> [...]

<pinc2>The terms of DeviantArt's privacy policy are incorporated into, and form a part of, these Terms.</pinc2>

Arbitration Clause

Is arbitration mandatory before the case can go to court?

If arbitration is fully optional: **clearly fair**

If arbitration should take place in a state other than the state of consumer's residence and/or be based on arbiter's discretion (i.e. not on law): **clearly unfair**

All other arbitration clauses: **potentially unfair**

Arbitration Clause

Is arbitration mandatory before the case can go to court?

A **clearly unfair** arbitration clause (Rovio):

<a3>Any dispute, controversy or claim arising out of or relating to this EULA or the breach, termination or validity thereof shall be finally settled at Rovio's discretion (i) at your domicile's competent courts; or (ii) by arbitration in accordance with the Rules for Expedited Arbitration of the Arbitration Institute of the Finland Chamber of Commerce. The arbitration shall be conducted in Helsinki, Finland, in the English language.</a3>

An example from the Instagram Terms of Service

We reserve the right, in our sole discretion, to change these Terms of Use ("Updated Terms") from time to time.

Unless we make a change for legal or administrative reasons, we will provide reasonable advance notice before the Updated Terms become effective. You agree that we may notify you of the Updated Terms by posting them on the Service, and that your use of the Service after the effective date of the Updated Terms (or engaging in such other conduct as we may reasonably specify) constitutes your agreement to the Updated Terms.



<lt2>TO THE EXTENT PERMITTED BY LAW, THE TOTAL LIABILITY OF GOOGLE, AND ITS SUPPLIERS AND DISTRIBUTORS, FOR ANY CLAIMS UNDER THESE TERMS, INCLUDING FOR ANY IMPLIED WARRANTIES, IS LIMITED TO THE AMOUNT YOU PAID US TO USE THE SERVICES (OR, IF WE CHOOSE, TO SUPPLYING YOU THE SERVICES AGAIN).</lt2>

<lt2>IN ALL CASES, GOOGLE, AND ITS SUPPLIERS AND DISTRIBUTORS, WILL NOT BE LIABLE FOR ANY LOSS OR DAMAGE THAT IS NOT REASONABLY FORESEEABLE.</lt2>

Business uses of our Services

If you are using our Services on behalf of a business, that business accepts these terms. It will hold harmless and indemnify Google and its affiliates, officers, agents, and employees from any claim, suit or action arising from or related to the use of the Services or violation of these terms, including any liability or expense arising from claims, losses, damages, suits, judgments, litigation costs and attorneys' fees.

About these Terms

<ch2>We may modify these terms or any additional terms that apply to a Service to, for example, reflect changes to the law or changes to our Services. You should look at the terms regularly.</ch2> We'll post notice of modifications to these terms on this page. We'll post notice of modified additional terms in the applicable Service. Changes will not apply retroactively and will become effective no sooner than fourteen days after they are posted. However, changes addressing new functions for a Service or changes made for legal reasons will be effective immediately. If you do not agree to the modified terms for a Service, you should discontinue your use of that Service.

If there is a conflict between these terms and the additional terms, the additional terms will control for that conflict.

These terms control the relationship between Google and you. They do not create any third party beneficiary rights.

If you do not comply with these terms, and we don't take action right away, this doesn't mean that we are giving up any rights that we may have (such as taking action in the future).

If it turns out that a particular term is not enforceable, this will not affect any other terms.

<law2>The laws of California, U.S.A., excluding California's conflict of laws rules, will apply to any disputes arising out of or relating to these terms or the Services.</law2> <j3>All claims arising out of or relating to these terms or the Services will be litigated exclusively in the federal or state courts of Santa Clara County, California, USA, and you and Google consent to personal jurisdiction in those courts.</j3>

The Machine Learning Methodology

From a ML point of view, we modelled the problem as:

a **detection task**: does a sentence contain a potentially unfair clause? Positive (if p unfair), Negative (otherwise)

a **sentence classification task**: to what category does the unfair clause belong?

Approaches

- Bag of Words (BoW): leverages on the lexical information in sentences
- Tree kernels: leverages on grammatical structure of sentences
- Convolutional Neural Networks, SVM, etc.

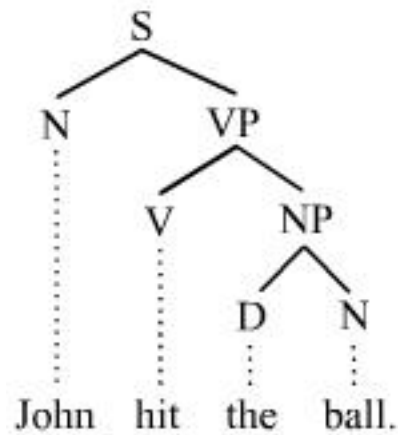
The Bag of Words Model (BoW)

- Build to leverage the lexical information in sentences
- Each word is a feature
- Each sentence is represented as a vector of features, as large as the dimension of vocabulary in the corpus (also bigrams)
- Each feature is either zero (if the word does not appear in the sentence), or different than zero (if it appears)
- We feed VECTORS to Support Vector Machines (SVMs)

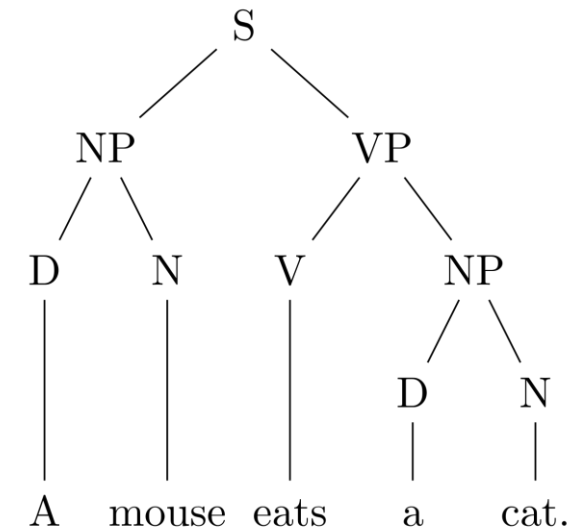
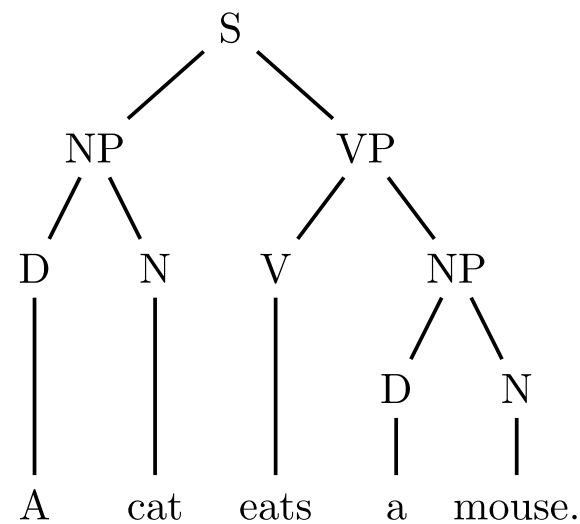
TEXT	VOCABULARY	VECTORS
<i>1. It was the best of times,</i>	It	[it, was, the, best, of, times, ...]
<i>2. it was the worst of times,</i>	was	1. [1, 1, 1, 1, 1, 1, 0, 0, 0, 0]
<i>3. it was the age of wisdom,</i>	the	2. [1, 1, 1, 0, 1, 1, 1, 0, 0, 0]
<i>4. it was the age of foolishness,</i>	Best	3. [1, 1, 1, 0, 1, 0, 0, 1, 1, 0]
	of	4. [1, 1, 1, 0, 1, 0, 0, 1, 0, 1]
	Times	
	Worst <- ONE OCCURRENCE	
	[...]	

The Tree Kernels Model

- A method for representing the **semantic** structure of sentences (Constituency-based parse tree)
- A method for comparing tree graphs to each other, allowing us to get quantifiable measurements of their similarities or differences
- A TK consists of a similarity measure between two trees, which takes into account the number of common substructures, known as fragments
- More sophisticated, lately proven to be effective in argumentation mining



Constituency-based parse tree

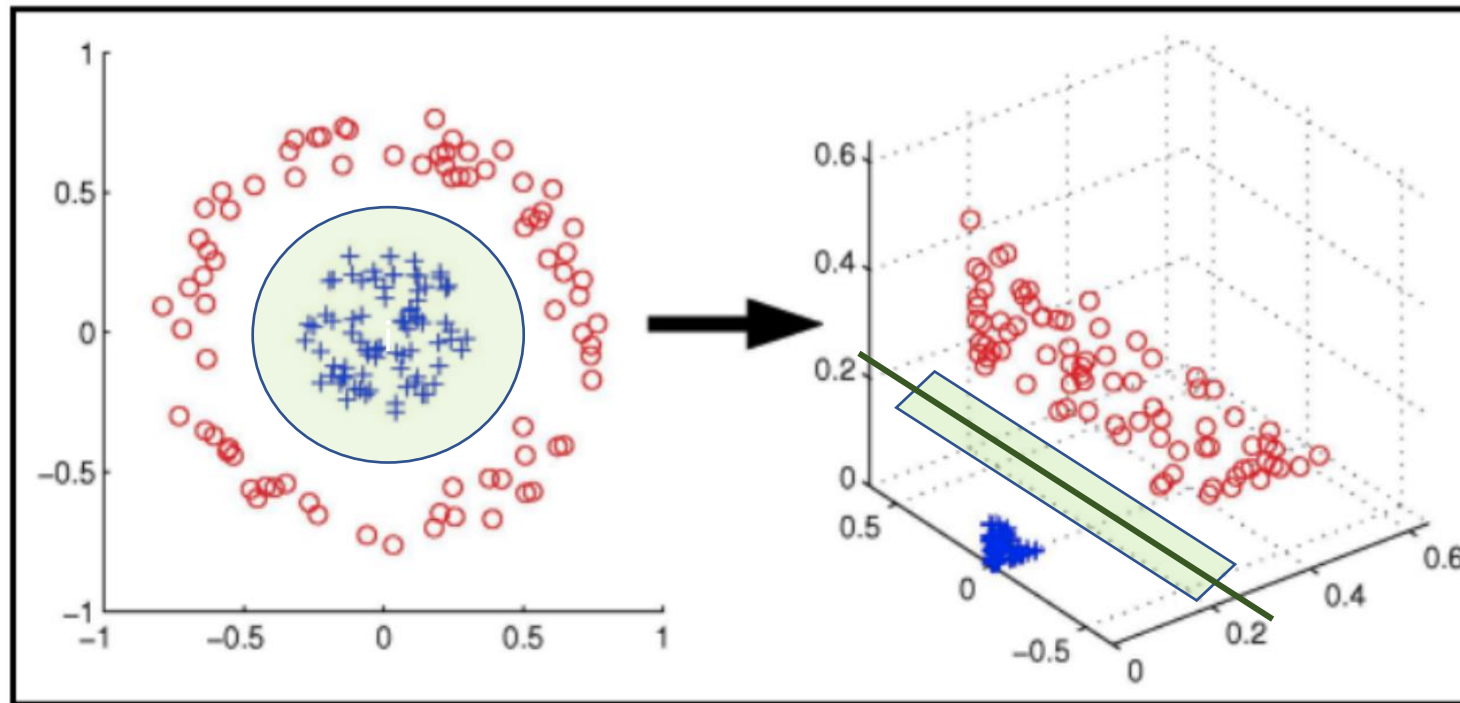


Supported Vector Machines (SVM)

Once we've defined our mathematical space (set of vectors mapping sentences) SVM can be used to detect whether a new clause is fair or not, and the category it belongs to.

SVM is a **binary classifier**. It builds an **hyper plane classifier** (in an augmented space).

SVMs calculate a **maximum-margin boundary** that leads to a homogeneous partition of all data points. This classifies a SVM as a **maximum margin classifier**.



Data representation and Ensemble Methods

The problem is formulated as a binary classification task where

- the positive class is either the union of all potentially unfair sentences
- or the set of potentially unfair clauses of a single category

Results of each configuration have been collected and compared to see which one performs better. The **better performance** is an **ensemble**.

C1: SVM exploiting BoW

C2: SVM exploiting TK for sentence representation

C3: SVM for collective classification of sentences in a document (BoW+TK)



Voting: if 2 out of 3 predict positive sentence

The input sentence is classified as potential unfair

Experiments

Leave-One-Out procedure: each document in turn, is used as test set, leaving the remaining documents for training set (4/5) and validation set (1/5) for model selection

3 Metrics

Precision: fraction of positive predictions, actually labelled as positive

Recall: fraction of positive examples that are correctly detected

F1: harmonic mean between precision and recall

Baselines for comparison: random classifier

Experimental Results

Performance: Training set size = 50 Tos

Method	P	R	F_1
SVM—single model	0.729	0.830	0.769
SVM—combined model	0.798	0.782	0.781
Tree kernels	0.777	0.718	0.739
Convolutional neural networks	0.729	0.739	0.722
Long short-term memory networks	0.696	0.723	0.698
SVM-HMM—single model	0.759	0.778	0.758
SVM-HMM—combined model	0.859	0.687	0.757
Ensemble (C1+C2+C3+C6+C7)	0.826	0.797	0.805
Random baseline	0.125	0.125	0.125
Always positive baseline	0.123	1.000	0.217

Best performing: **Ensemble**

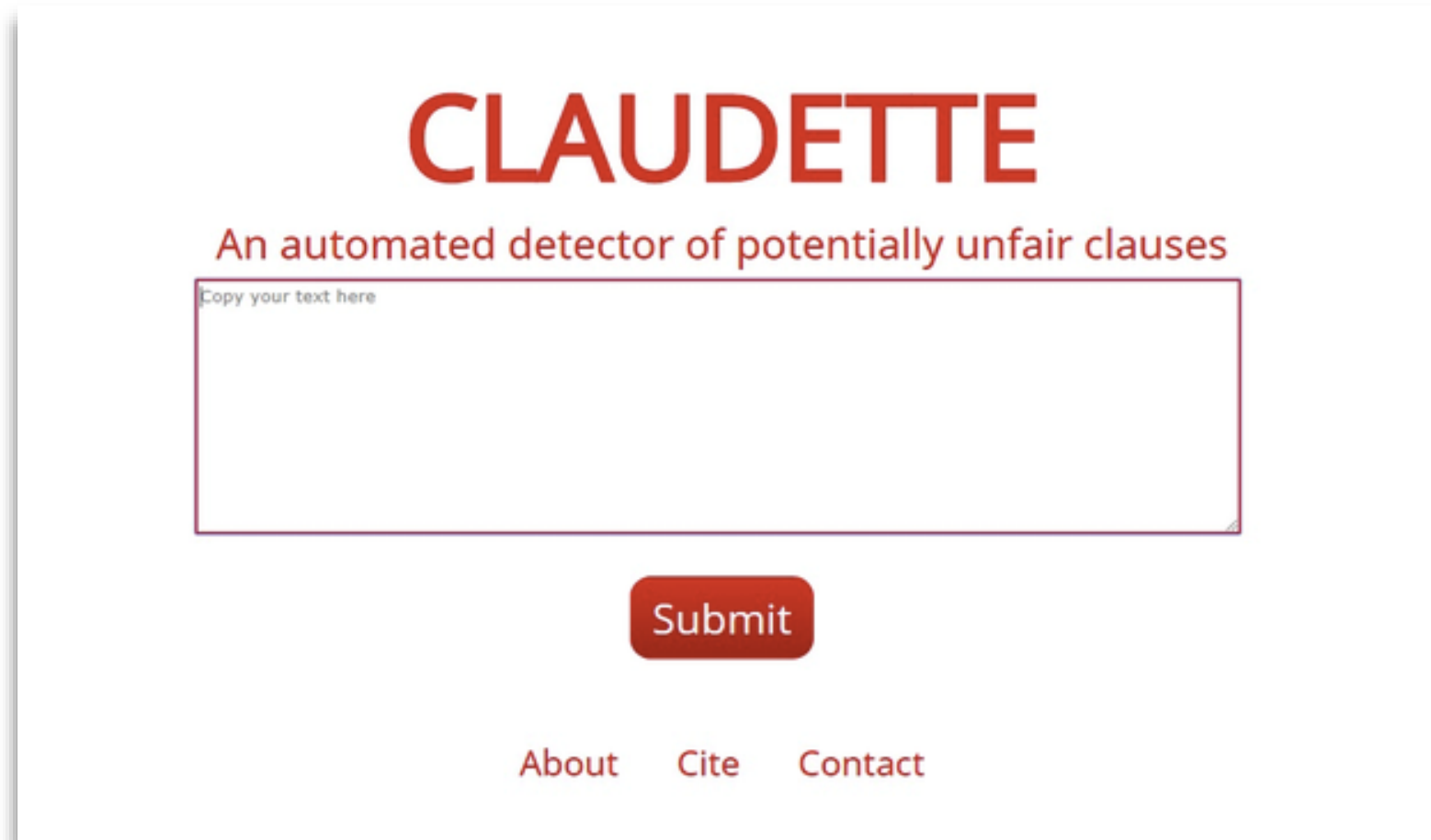
claudette.eui.eu

Experimental Results

Claudette correctly detected around **80%** of the **potentially unfair** clauses in each category, ranging from a **minimum 72.7%** in the case of arbitration clauses, **up to 89.7%**, as in the case of jurisdiction clauses.

Tag	Precision	Recall	F ₁
Arbitration	0.832	0.814	0.823
Unilateral change	0.832	0.814	0.823
Content removal	0.713	0.780	0.745
Jurisdiction	1.000	0.941	0.970
Choice of law	0.984	0.886	0.932
Limitation of liability	0.961	0.905	0.932
Unilateral termination	0.786	0.932	0.853
Contract by using	0.949	0.957	0.953

An online server



The screenshot shows the CLAUDETTE website interface. At the top, the word "CLAUDETTE" is written in large, bold, red capital letters. Below it, the subtitle "An automated detector of potentially unfair clauses" is displayed in a smaller red font. A large, empty text input box with a red border is centered on the page, containing the placeholder text "Copy your text here" in a small, light gray font. Below the input box is a prominent red button with rounded corners and the word "Submit" in white text. At the bottom of the interface, three links are listed in red text: "About", "Cite", and "Contact".

CLAUDETTE

An Automated Detector of Potentially Unfair Clauses

Claudette found 3 potentially unfair clauses (displayed in **bold**) out of 295 sentences.
Below you can find a summary of the detected clauses, possibly linked to the most plausible rationales.

Potentially unfair clause #1

EXCEPT FOR CERTAIN TYPES OF DISPUTES MENTIONED IN THE ARBITRATION CLAUSE , YOU AND HEADSPACE AGREE THAT DISPUTES RELATING TO THESE TERMS OR YOUR USE OF THE PRODUCTS WILL BERESOLVED BY MANDATORY BINDING ARBITRATION , AND YOU WAIVE ANY RIGHT TO PARTICIPATE IN A CLASS-ACTION LAWSUIT OR CLASS-WIDE ARBITRATION .

Unfairness categories: **Arbitration**

[Hide/show rationales](#)

Potentially unfair clause #2

1.4 CHANGES TO TERMS Headspace reserves the right to change or update these Terms , or any other of our policies or practices , at any time , and will notify users by posting such changed or updated Terms on this page .

Unfairness categories: **Unilateral Change**

[Hide/show rationales](#)

Potentially unfair clause #3

Your continued use of the Products constitutes your agreement to abide by the Terms as changed .

Unfairness categories: **Contract by Using**

[Hide/show rationales](#)

CLAUDETTE

An Automated Detector of Potentially Unfair Clauses

Potentially unfair clause #1

EXCEPT FOR CERTAIN TYPES OF DISPUTES MENTIONED IN THE ARBITRATION CLAUSE , YOU AND HEADSPACE AGREE THAT DISPUTES RELATING TO THESE TERMS OR YOUR USE OF THE PRODUCTS WILL BERESOLVED BY MANDATORY BINDING ARBITRATION , AND YOU WAIVE ANY RIGHT TO PARTICIPATE IN A CLASS-ACTION LAWSUIT OR CLASS-WIDE ARBITRATION .

Unfairness categories: **Arbitration**

[Hide/show rationales](#)

Potentially unfair clause #2

1.4 CHANGES TO TERMS Headspace reserves the right to change or update these Terms , or any other of our policies or practices , at any time , and will notify users by posting such changed or updated Terms on this page .

Unfairness categories: **Unilateral Change**

[Hide/show rationales](#)

The clause is potentially unfair for **Unilateral Change** since the provider has the right for unilateral change of the contract, services, goods, features for any reason at its full discretion, at any time (score = 0.834)

Potentially unfair clause #3

Your continued use of the Products constitutes your agreement to abide by the Terms as changed .

Unfairness categories: **Contract by Using**

[Hide/show rationales](#)



Human Legal experts are able to recognize potentially unfair clauses thanks to their **background knowledge** of the domain.

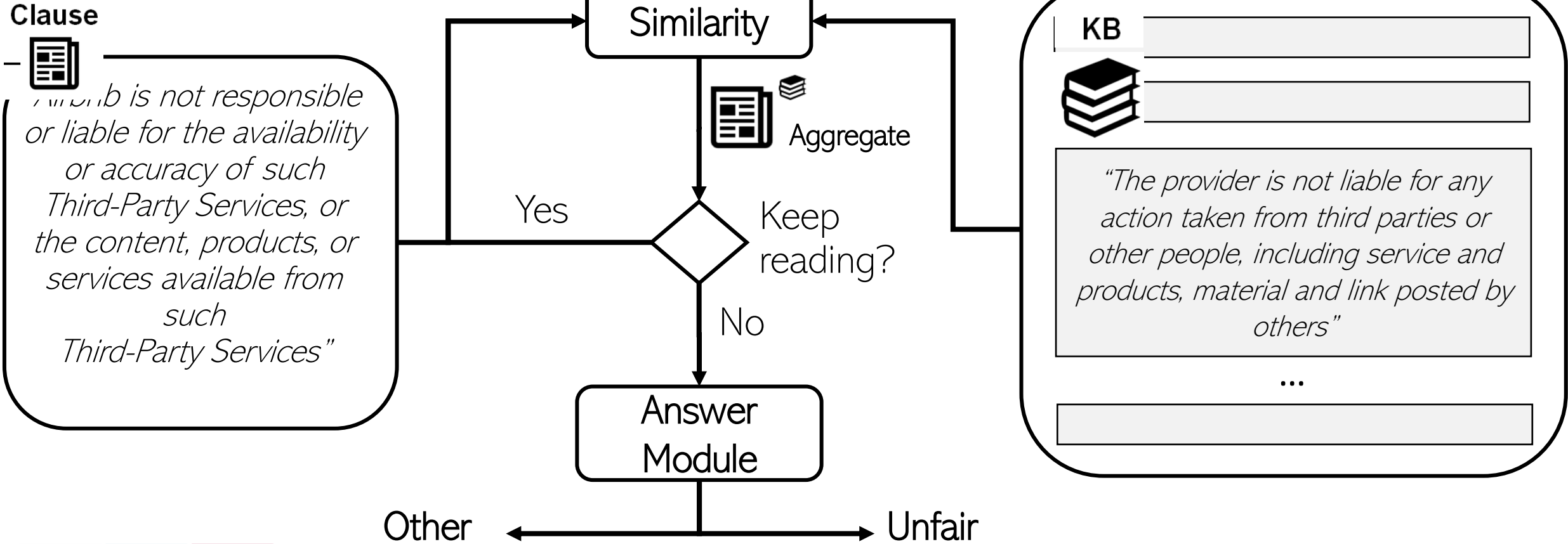
- Rely on intuitions, trained on experience with relevant examples
- Able to explain their intuitions of unfairness, provide reasons why a clause is unfair (**Legal Rationales**), and use rationales to guide such intuitions
- Appealing to their background knowledge (e.g. Standards, Rules and Principles, Judicial precedents) as support for reasoning

Memory-Augmented Neural Networks

- Process input and **store** the information in some **memory**
- Understand **pieces of knowledge** relevant to a given **query**
- Retrieve **concepts** from memory
- Combine **memory and query** to make a prediction

Exploiting Knowledge for Unfairness Identification

Esperimental Setup



CLAUDETTE meets GDPR

Moving to **privacy policies**: **what is different?**

Ensure some information is **present** and **complete** (i.e., compliant with articles 13-14 of GDPR)

Detect problematic clauses for **data processing**

Detect **vague** language

CLAUDETTE meets GDPR

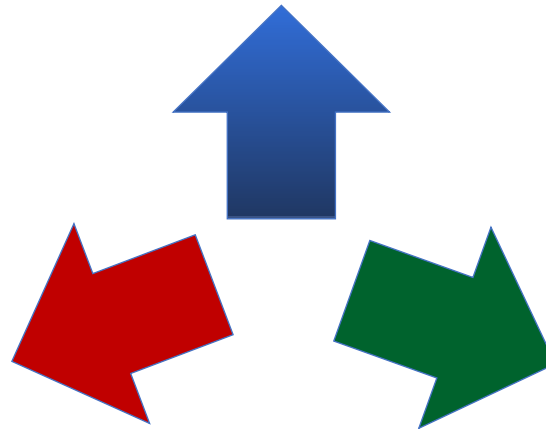
The Golden Standard: Lawfulness Fairness Transparency

Comprehensiveness of information

The policy should contain all the information required by articles 13 and 14 of the GDPR.

Clarity of expression

The policy should be framed in an understandable and precise language.



Substantive compliance

The policy should only allow for processings of personal data that are compliant with the GDPR.

Different Levels of Achievement: Optimal and Suboptimal (questionable or insufficient)

Comprehensiveness of information

23 categories (GDPR art 13 and 14)

Type of required information	Symbol
Identity of the controller (controller's representative)	<id>
Contact details of the controller (controller's representative)	<contact>
Contact details of the data protection officer	<dpo>
Purposes of the processing	<purp>
Legal Basis for the processing	<basis>
Categories of personal data concerned	<cat>
Recipients or categories of recipients of the personal data	<recep>
Period for which the personal data will be stored, or the criteria used to determine that period	<ret>
Right to lodge a complaint with a supervisory authority	<complain>
...	<...>

Purposes of the processing for which the personal data are intended

13(1)(c) and 14(1)(c))

Clauses where the purposes of the processing are exhaustive and not vague: **fully informative**

In other cases (e.g. when a clause only provides examples): **insufficiently informative**

WhatsApp Privacy Policy (last updated on 24 April 2018)

```
<purp2> WhatsApp must receive or collect some information to operate, provide, improve, understand, customize, support, and market our Services, including when you install, access, or use our Services.</purp2>
```



The Categories of personal data concerned

Clauses where the categories of personal data are comprehensively specified and not vague: **fully informative**

In other cases (e.g. when a clause only provides examples): **insufficiently informative**

Google Privacy Policy (last updated on 25 May 2018)

```
<cat1>We collect information about your location when you use our services, which helps us offer features like driving directions for your weekend getaway or showtimes for movies playing near you.</cat1>
```

The Categories of personal data concerned

Clauses where the categories of personal data are comprehensively specified and not vague: **fully informative**

In other cases (e.g. when a clause only provides examples): **insufficiently informative**

Edreams Privacy Policy (last updated on 25 May 2018)

```
<cat2>If you sign up for our website using your social media account, link your account on our website to your social media account, or use certain other social media features of ours, we may access information about you via that social media provider in accordance with the provider's policies.</cat2>
```

Substantive compliance

10 categories (GDPR art 5, 6, 9 and others)

Type of clause	Symbol
Processing of special categories of personal data (e.g. health, sex life, political opinions, religious beliefs, etc.)	<sens>
Consent by using	<cuse>
Take or leave it approach	<tol>
Third party data transfers	<tp>
Policy change	<pch>
Transfer of data to third countries	<cross>
Processing of children's data	<child>
Licensing data	<lic>
Advertising	<ad>
Any other type of consent	<c>

Consent by using

(GDPR art 4(11)) (Rec 32)

When the consent is explicitly required: **fair processing clause**

Clauses stating that by simply using the service, the user consents to the terms of the privacy policy: **unfair processing clauses**

Epic games Privacy Policy (last updated on 24 May 2018)

<cuse3> when you use our websites, games, game engines, and applications, you agree to our collection, use, disclosure, and transfer of information as described in this policy, so please review it carefully.**</cuse3>**



Policy change

When notice is given and new consent is required: **fair processing clause**

When notice is given but a new consent (or confirmation of reading) is not required: **problematic processing clause**

Twitter Privacy Policy (effective on 25 May 2018)

<pch2>We may revise this Privacy Policy from time to time. The most current version of the policy will govern our processing of your personal data and will always be at <https://twitter.com/privacy>. If we make a change to this policy that, in our sole discretion, is material, we will notify you via an @Twitter update or email to the email address associated with your account.**</pch2>**

Policy change

When no notice is given and new consent is not required: **unfair processing clause**

Booking Privacy Policy (last updated on 9 May 2018)

<pch3>We might amend the Privacy Statement from time to time. If you care about your privacy, visit this page regularly and you'll know exactly where you stand.</pch3>

Clarity of expression

(GDPR art. 5(1)(a), 12(1) and others)

Is the privacy policy framed in an understandable and precise language?

4 main indicators of vagueness

Indicator	Language qualifiers
1. Conditional Terms The performance of a stated action or activity is dependent on a variable trigger	Depending, as necessary, as appropriate, as needed, otherwise reasonably, sometimes, from time to time, etc.
Example	Rationale
<vag> We also may share your information if we believe, in our sole discretion, that such disclosure is necessary :... "</vag>	The practice described as “necessary” suggests that the sharing will only occur in exceptional cases, however the clause fails to specify under what exceptional conditions the provider will disclose the information.

Clarity of expression

Indicator	Language qualifiers
2. Generalization: i.e. terms that vaguely abstract information practices using contexts that are unclear. Action(s)/Information Types are vaguely abstracted with unclear conditions.	generally, mostly, widely, general, commonly, usually, normally, typically, largely, often, primarily, among other things, etc.
Example	Rationale
<pre><vag> We typically or generally collect information ...</vag> <vag> When you use an Application on a Device, we will collect and use information about you in generally similar ways and for similar purposes as when you use the TripAdvisor website.</vag></pre>	The use of the generalization term “generally” obscures for the data subject the service provider activities, since it provides a large flexibility to the service provider.

Clarity of expression

(GDPR art. 5(1)(a), 12(1) and others)

Indicator	Language qualifiers
<p>3. Modality: it includes modal verbs, adverbs and non-specific adjectives, which create uncertainty with respect to actual action; it includes whether an action is possible. Modality does not include whether an action and/or activity is permitted. Modality mainly refers to ambiguous possibility of action or event.</p>	<p>may, might, could, would, possible, possibly, etc.</p>
Example	Rationale
<p><vag>We may use your personal data to develop new services </vag></p>	<p>it is unclear whether or not the controller will use the data subject information to develop new services and in what cases and under</p>

Clarity of expression

Indicator	Language qualifiers
4. Non specific Numeric quantifiers: which create ambiguity as to the actual measure	certain, numerous, some, most, many, various, including (but not limited to), variety
Example	Rationale
<code><vag>When you create an Apple ID, apply for commercial credit, purchase a product, download a software update, register for a class at an Apple Retail Store, connect to our services, contact us or participate in an online survey, we may collect a variety of information, including your name, mailing address, phone number, email address, contact preferences, device identifiers, IP address, location information and credit card information.</vag></code>	it creates ambiguity with regard to the actual measure of information the data controller collect

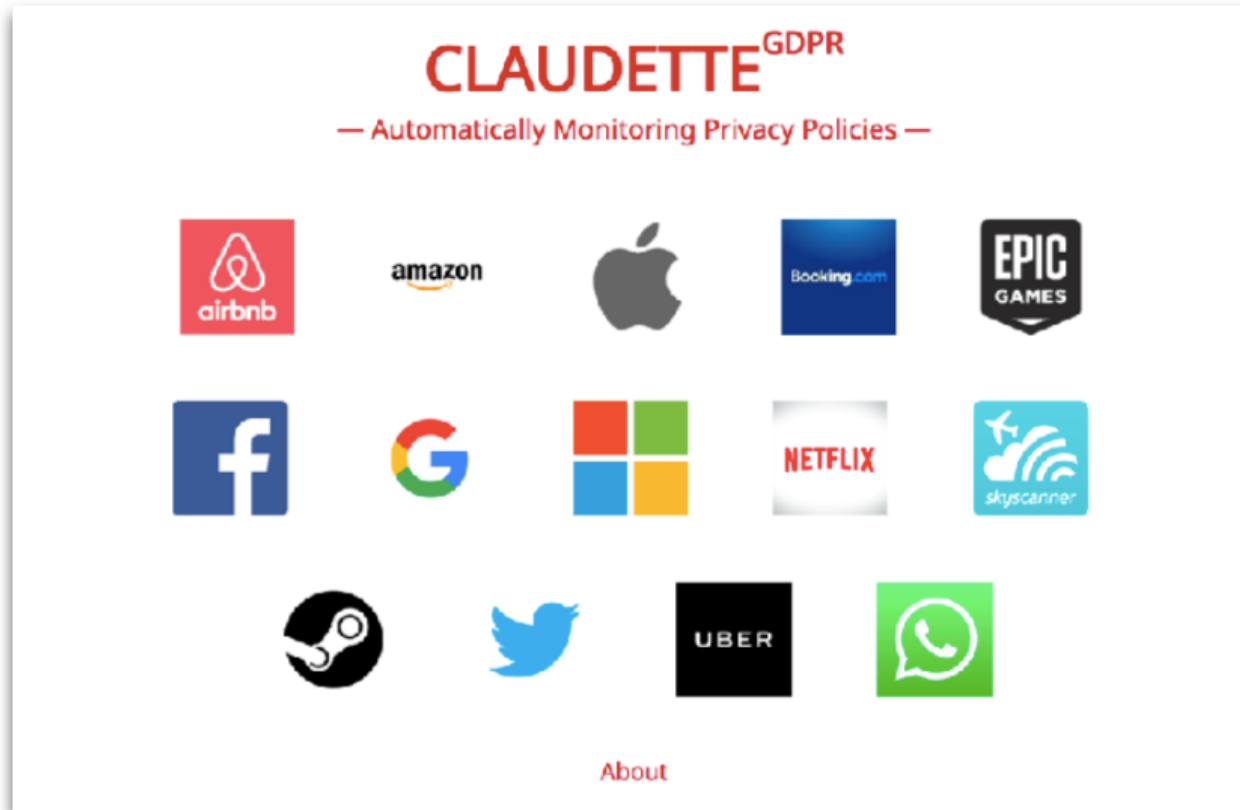
Clarity of expression

A combination of different forms of vagueness

<vag>We **generally may** share personal information we collect with **certain** service providers, **some** of whom **may** use information for their own purposes **as necessary**.</vag>

CLAUDETTE FOR GDPR

<http://claudette-gdpr.eu>



Manually annotated corpus

WHERE DID WE START?

- 14 documents (32 now)
- 3,658 sentences
- 80,398 words
- 11.0% sentences contain unclear language
- 33.9% sentences contain potentially unlawful clauses

CLAUDETTE FOR GDPR

<http://claudette-gdpr.eu>

CLAUDETTE^{GDPR}
— Automatically Monitoring Privacy Policies —

Use the checkboxes below to choose the sentences you want to highlight.
In parenthesis, you can see the number of detected sentences for each category.
By hovering your cursor over each unfair sentence, you can see the categories of each detected sentence.
If you want to see the entire policy, click the grey [...] symbol; otherwise you will see just the detected clauses.

Apple Privacy Policy
(153 sentences, 3643 words, 20861 characters)

Full Information (6) Insufficient Information (19) Unclear Language (12) Problematic Processing (24)

[...]

You may be asked to provide your personal information anytime you are in contact with Apple or an Apple affiliated company .
Apple and its affiliates may share this personal information with each other and use it consistent with this Privacy Policy .
They may also combine it with other information to provide and
You are not required to provide the personal information that we use to create, develop, operate, deliver, and improve our products, services, content and advertising, and for loss prevention and anti-fraud purposes .

Insufficient information clause for categories:
•Categories of personal data concerned

Unclear language clause

[...]

When you share your content with family and friends using Apple products , send gift certificates and products , or invite others to participate in Apple services or forums , Apple may collect the information you provide about those people such as name , mailing address , email address , and phone number .
Apple will use such information to fulfill your requests , provide the relevant product or service , or for anti-fraud purposes .
In certain jurisdictions , we may ask for a government issued ID in limited circumstances including when setting up a wireless account and activating your device , for the purpose of extending commercial credit , managing reservations , or as required by law .

[...]

The personal information we collect allows us to keep you posted on Apple 's latest product announcements , software updates , and upcoming events .
If you do not want to be on our mailing list , you can opt out anytime by updating your preferences .
We also use personal information to help us create , develop , operate , deliver , and improve our products , services , content and advertising , and for loss prevention and anti-fraud purposes .

Automated tagging prototype

<http://155.185.228.137/claurette4gdpr/>

CLAUDETTE^{GDPR}

Detecting Unclear Language in Privacy Policies (beta version)

Copy your text here

Submit

[About](#) [Cite](#) [Contact](#)

Accuracy: 68%

Automated tagging prototype

<http://155.185.228.137/claurette4gdpr/>

CLAUDETTE^{GDPR}

Detecting Unclear Language in Privacy Policies (beta version)

Claurette found 56 clauses with unclear language (displayed in **bold**) out of 229 sentences.

[...]

Learn more: Personal information we collect **We collect personal information from you and any devices (including mobile devices) you use when you: use our Services, register for an account with us, provide us information on a web form, update or add information to your account, participate in our promotions, or communicate with us via email, text message, or discussion chat, or when you otherwise correspond with us.**

Unclear Language unfair clause

Some of this personal information, such as a way to identify you, is necessary to enter into our User Agreement.

The provision of all other personal information is voluntary, but may be necessary in order to use our Services, such as the bidding, buying or selling information needed to conclude a transaction.

We may also collect personal information from other sources, as described below.

Personal information you give us when you use our Services or register for an account with us Identifying information such as your name, addresses, telephone numbers or email addresses when you register for an account with us Bidding, buying, or selling information you provide during a transaction, or other transaction-based content that you generate or that is connected to your account as a result of a transaction you are involved in Other content that you generate, or that is connected to your account (such as adding items to your basket, adding items to your Watch List, creating collections, and following other collections and sellers) Financial information (such as credit card or bank account numbers) in connection

Multilingualism: the German, Italian and Polish Claudette for ToS and PPs

 **CLAUDETTE** 

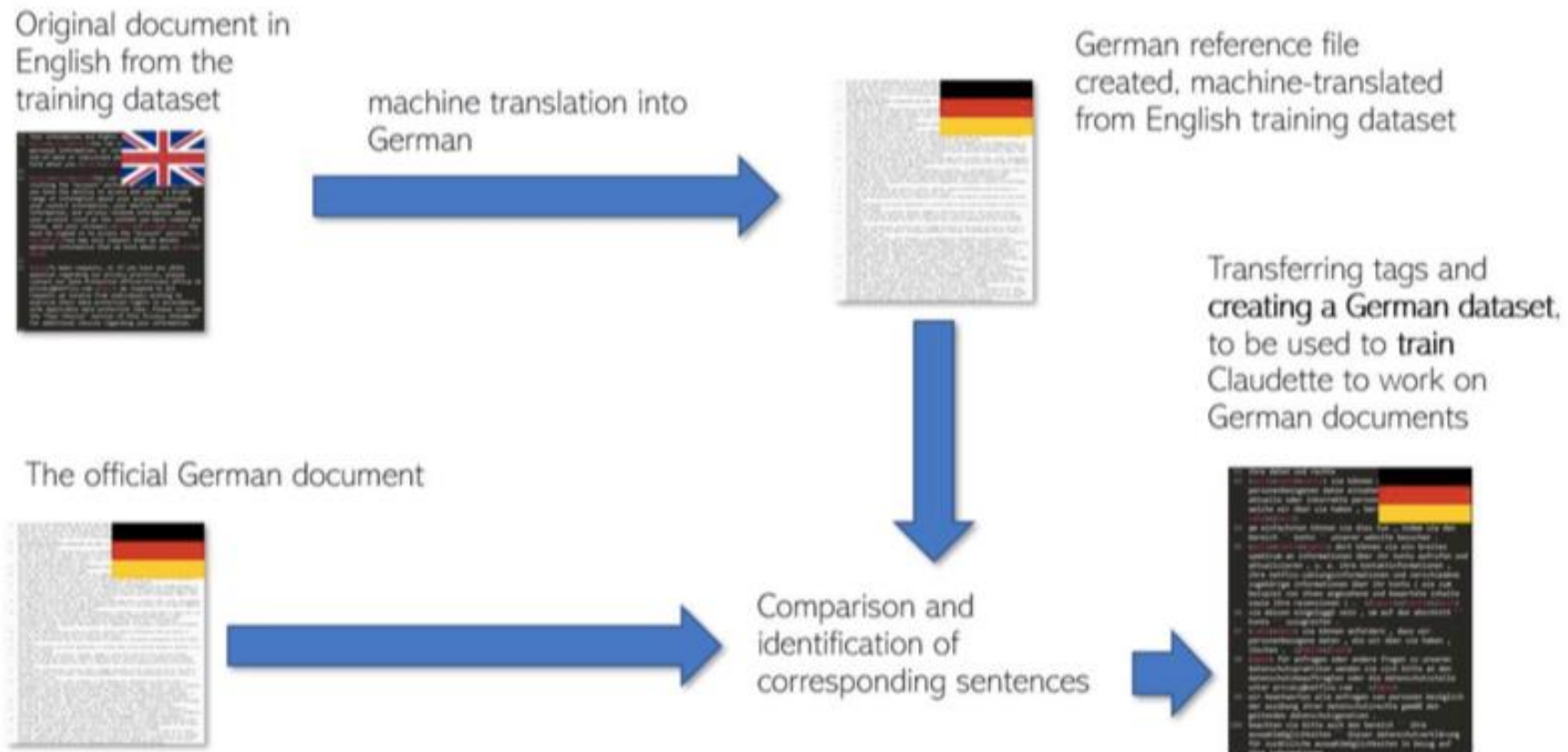
Ein automatisierter Detektor für potenziell unlautere Klauseln
(GERMAN BETA VERSION)

Kopieren Sie Ihren Text hier

Submit

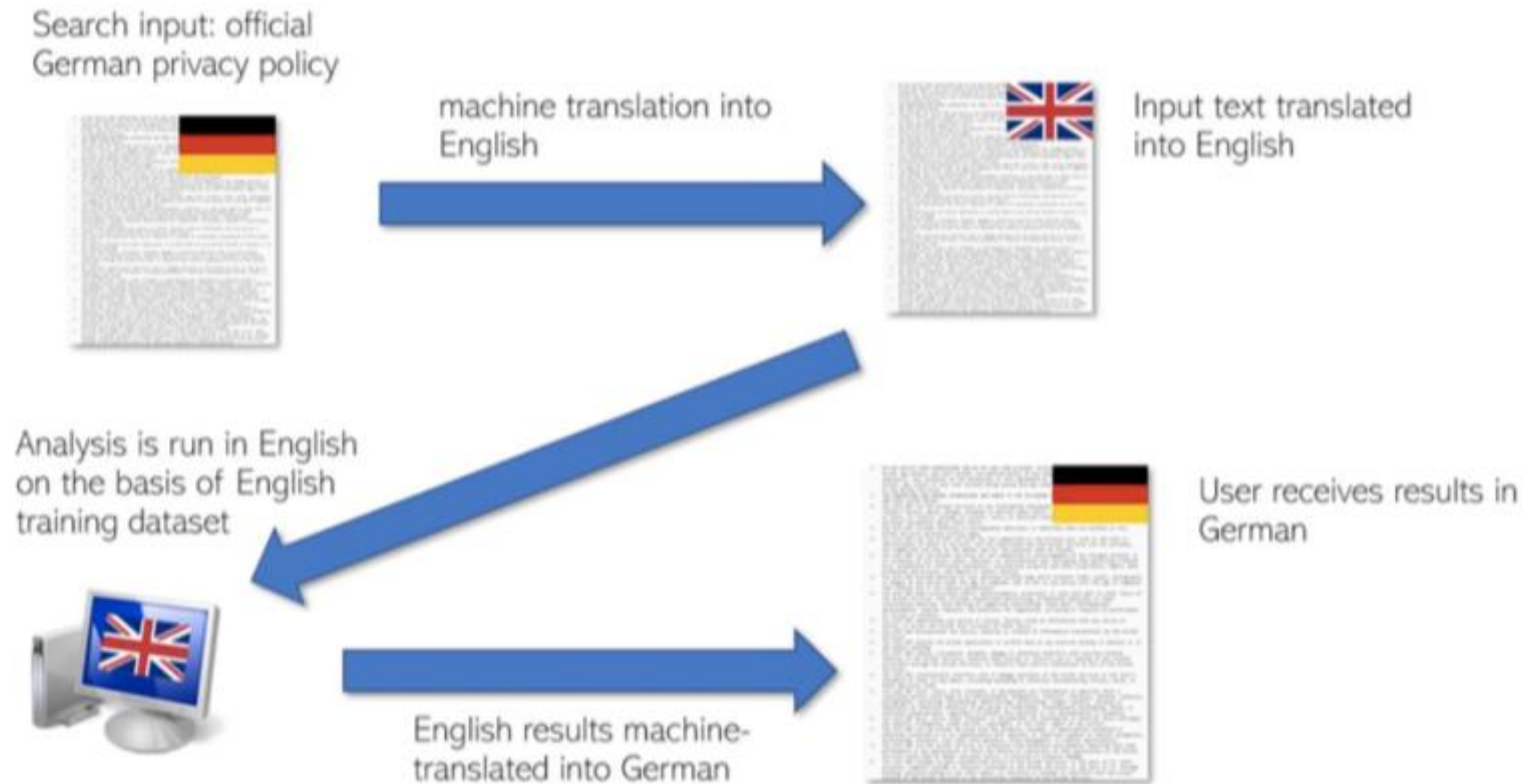
Multilingualism: the German, Polish and Italian Claudette

Approach 1: Semi-automated creation of a target-language dataset



Multilingualism: the German Claudette

Approach 2 – Machine translation of queries



Multilingualism: the German Polish Italian Claudette

Approach 1 – Semi-automated creation of a target-language dataset

	DE	EN	IT	PL
a1	3	3	4	4
a2	22	29	29	35
a3	3	4	4	4
ch2	98	100	103	103
ch3	1	1	0	0
cr2	25	27	28	26
cr3	23	24	24	26
j1	14	15	15	15
j3	46	49	48	50
law1	18	16	16	19
law2	33	39	36	36
ltd1	19	27	16	17
ltd2	212	229	216	229
ltd3	1	1	1	1
pinc2	17	21	20	21
ter2	69	71	71	75
ter3	49	49	50	49
use2	54	58	58	61
total	707	753	739	771

Dataset: multilingual parallel corpus consisting of 25 Terms of Service annotated in English, Italian, German and Polish.

Table 1: Corpus statistics. we report the number of annotated clauses for each tag, across the four different languages. Suffices 1, 2, and 3 represent levels of fairness: 1 means clearly fair, 2 stays for potentially unfair, and finally 3 for clearly unfair.

claudette.eui.eu

Multilingualism: the German Polish Italian Claudette

Approach 1 – Semi-automated creation of a target-language dataset

	DE	IT	PL
Precision	0.87	0.94	0.90
Recall	0.95	0.98	0.97
F1-macro	0.91	0.94	0.91
F1-micro	0.91	0.96	0.93
F1-weighted	0.91	0.96	0.93

Table 2: Projection results for the three languages.

WEB-CRAWLER

Developed as a tool for automatic privacy policy monitoring

Two types of monitoring:

- Checking the date on the document
- Comparison of the content with the previously saved version

Earnings reports by e-mail

Selected Publications

- LRuggeri F., Lagioia F., Lippi M., Torroni P., (2021) Detecting and explaining unfairness in consumer contracts through memory networks, in *Artificial Intelligence and Law*, Springer- Nature, 1-34
- Lippi, M.; Contissa, G.; Jablonowska, A.; Lagioia, F.; Micklitz, H-W; Palka, P.; Sartor, G.; Torroni, P., *The Force Awakens: Artificial Intelligence for Consumer Law*, *The journal of Artificial Intelligence Research*, 2020, 67,, 169 - 190
- Lippi, M., Palka, P., Contissa, G., Lagioia, F., Micklitz, H. W., Sartor, G., & Torroni, P. CLAUDETTE: an Automated Detector of Potentially Unfair Clauses in Online Terms of Service, *Artificial Intelligence and Law*, Springer (2019).
- Lippi, M., Contissa, G., Lagioia, F., Micklitz, H. W., Palka, P., Sartor, G., & Torroni, P., Consumer protection requires artificial intelligence. *Nature Machine Intelligence*, 1, (2019).
- Contissa, G., Docter, K., Lagioia, F., Lippi, M., Micklitz, H.-W., Palka, P., Sartor, G., Torroni, P., CLAUDETTE meets GDPR: Automating the Evaluation of Privacy Policies using Artificial Intelligence Study Report, (BEUC), 2018.
- Lippi, M., Palka, P., Contissa, G., Lagioia, F., Micklitz, H. W., Sartor, G., & Torroni, P., Towards Consumer-Empowering Artificial Intelligence, JCAI-ECAI, Stockholm, special track on the evolution of contours of AI, (2018) .

MAI4CAREU

Master programmes in Artificial
Intelligence 4 Careers in Europe

